20040804 110

Detection of a PSK Signal Transmitted Through a Hard-Limited Channel

PRAVIN C. JAIN AND NELSON M. BLACHMAN

Abstract—This paper considers the problem of the detection of a binary phase-modulated carrier which has been transmitted along with noise through a hard-limiting repeater, corrupted by additional noise, and demodulated by cross correlation and sampling at the receiver. Three equivalent expressions are obtained for the error probability of the receiver output. Two of these expressions take the form of an infinite series involving either confluent hypergeometric functions or modified Bessel functions. A third form allows representation of the error probability in terms of Rice's Ie function. In the special case where the signal-to-noise ratios $\rho_1{}^2$ and $\rho_2{}^2$ at the limiter and receiver inputs are equal, the error probability is simply

$$P_e = \tfrac{1}{2} \exp(-\rho_1{}^2).$$

Asymptotic expressions for the error probability for large and small signal-to-noise ratios are also derived. The error probability is found to be smaller than that of a linear system for all practical values of limiter and receiver input signal-to-noise ratios. The optimum repeater nonlinearity is investigated and is shown to be a limiter at such signal-to-noise ratios.

I. INTRODUCTION

THE MODEL for the communication system to be considered in this paper is shown in Fig. 1. A binary phase-shift-keyed (PSK) signal is assumed to get from the transmitter to the receiver via a hard-limiting repeater in the transmission channel. Additive Gaussian noise is introduced on both the uplink (repeater noise) and the downlink (receiver noise). Detection of the signal at the receiver is accomplished by a cross correlation-and-sampling operation. The receiver is assumed to heterodyne the received signal with a correctly phased local reference and to base its decision on the sign of the baseband output of a zonal low-pass filter that is sampled once during each data bit.

The results for this single-sample detection model can be extended to include the effect of postdetection integration over the full bit duration which is customary in correlation detection of binary PSK signals, via the commonly assumed multiple sampling and majority decision [1]–[3]. This is achieved by approximating the integration operation by a sum of TW independent receiver–output samples taken at
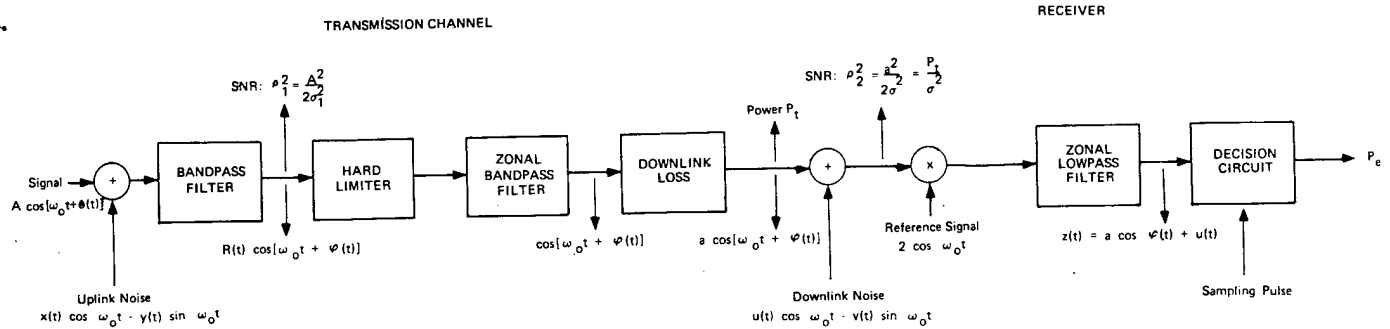
Fig. 1. Block diagram of PSK system.

the Nyquist rate and supposing that on the basis of these samples, TW independent decisions are made on each bit; a final overall decision is then taken on that bit by a majority vote. It is preferable that TW—the product of the bit duration times the signal bandwidth—be odd. The probability of error P_E is then equal to the probability that more than half the decisions will be in error and is given by the binomial distribution

$$P_E = \sum_{k=0}^{(TW-1)/2} \binom{TW}{k} (1 - P_e)^k P_e^{TW-k} \qquad (1)$$

where P_e is the probability of error for any one sample. Jacobs [1] has plotted P_E as a function of $TW[1 - P_e]^2$ for various values of TW. In this paper we determine P_e on the understanding that the error probability with post-detection integration can then be calculated from (1) or by using the plots given by Jacobs. Since P_E is a monotonic function of P_e, the conclusions reached in this paper concerning the helpfulness of limiting with single-sample detection also apply with postdetection integration. This is in agreement with the results and conclusions of Davisson and Milstein [2] and Lyons [3].

The calculation of the error rate with a hard-limiting repeater is of considerable interest both because limiters are often used at the inputs to traveling-wave-tube amplifiers when maximum output power is demanded of them and because at reasonably high signal-to-noise ratios, as we shall show, for a given average repeater output power, the error probability is smaller with a limiting repeater than with a linear one. We shall also see that a limiter is the optimum repeater nonlinearity at high signal-to-noise ratios. Since this superiority of the limiter seems to strike some communication theorists as incredible, we begin with some heuristic comments on this matter.

A. The Desirability of Limiting

Although cross correlation detection of binary PSK signals in noise is optimum, it is optimum only among the things that can be done solely at the receiver. In the present context optimization would have to include the best signal processing by the repeater for a given average repeater output power. Because keying simply switches the carrier phase between 0 and 180°, the repeater should waste no power on the quadrature component of the noise and should

put all of its power into retransmitting the inphase component of the signal-plus-noise, which is a sufficient statistic in regard to mark-versus-space decisions. The optimum repeater will therefore cross correlate, decide, and regenerate the signal corresponding to its decision.

It is easy to see that such a signal-regenerating repeater will yield performance superior to that of a linear repeater. For example, if the up and downlink signal-to-noise ratios are both reasonably high and are equal, the net overall signal-to-noise ratio at the receiver with a linear repeater will be 3 dB less, making the error probability many times that of a regenerating repeater, where the error probability is merely doubled by the two decisions in tandem. In fact, if we compare a chain of linear repeaters with a chain of regenerating repeaters along the path from the transmitter to the receiver with the product of their signal-plus-noise-to-signal ratios held constant (being equal to the overall net signal-plus-noise-to-signal ratio for the linear chain), we see that the error probability remains fixed for the linear system as the number of repeaters is increased. However, it approaches zero for the regenerating repeaters, as the signal-to-noise ratio for each repeater becomes proportional to the number of repeaters, thus reducing the error probability for each to a very small value—going to zero even when multiplied by the number of repeaters.

At high signal-to-noise ratios the inphase output component of a limiter will be very nearly the same as that of a regenerating repeater, but it is not possible to get a vanishing error probability with a chain of limiters, since the phase errors of their outputs (due to the quadrature noise) will accumulate in exactly the same way that the inphase component of the noise builds up in a chain of linear repeaters. An error results with limiters when the quadrature noise components take the signal phasor a quarter of the way or more around a circle about the origin, whereas in the linear case an error occurs when the inphase noise totals more than the length of the signal phasor and is opposite in sign. Since a 90° arc is $\pi/2$ times as long as the radius, limiting provides an improvement of nearly $20 \log_{10} \pi/2 = 3.82$ dB over a linear system—"nearly" because a 90° phase change in *either* direction will produce an error. Having established the plausibility of an error probability reduction due to limiting at high signal-to-noise ratios, we now proceed to treat this topic more thoroughly by means of a detailed mathematical analysis.

In Section II we will obtain first an expression for the output of the zonal low-pass filter in Fig. 1 and then derive three equivalent expressions for the error probability including asymptotic results for large and small signal-to-noise ratios. Numerical results for the error probability are provided in Section III, and the performance of a hard-limited system is compared with that of a linear repeater system. The investigation of the optimum repeater nonlinearity will be the topic of Section IV, with conclusions provided in Section V.

II. MATHEMATICAL ANALYSIS

The communication system shown in Fig. 1 has been recently investigated by Jain [4] and Davisson and Milstein [2]. Our treatment in this paper will simplify and extend some of the derivations and results obtained in [4] and [2], as well as many new results including entirely new material on the optimization of the nonlinearity.

A. Calculation of Receiver Output

The input to the limiter in Fig. 1 consists of a binary phase-modulated signal of frequency f_0 and amplitude A plus zero-mean, stationary Gaussian noise of rms value σ_1. The bandpass filter preceding the limiter is assumed to be wide enough to pass the signal with negligible distortion and to limit the uplink noise to a bandwidth that is small compared with the center frequency of the filter. The limiter input may be expressed as

$$f_i(t) = A \cos \left[\omega_0 t + \theta(t) \right] + x(t) \cos \omega_0 t - y(t) \sin \omega_0 t$$

$$= R(t) \cos \left[\omega_0 t + \phi(t) \right] \tag{2}$$

where the envelope $R(t)$ and the phase $\phi(t)$ are given by

$$R(t) = \sqrt{\left[A \cos \theta(t) + x(t) \right]^2 + y^2(t)}$$

$$\phi(t) = \arctan \frac{y(t)}{A \cos \theta(t) + x(t)}. \tag{3}$$

During any bit interval, $\theta(t)$ is either 0 or π, depending upon whether a mark or a space is being transmitted.

The bandpass limiter is assumed to be ideal in the sense that its fundamental-zone output $f_0(t)$ is given by

$$f_0(t) = \cos \left[\omega_0 t + \phi(t) \right] \tag{4}$$

i.e., envelope variation has been completely removed without distorting the phase modulation.

The signal is then transmitted to the receiver and noise is added to it on the downlink in the receiver front end. We assume this noise is stationary, zero-mean, Gaussian of rms value σ, and is independent of the uplink noise. Hence, the receiver input may be expressed as

$$g(t) = a \cos \left[\omega_0 t + \phi(t) \right] + u(t) \cos \omega_0 t - v(t) \sin \omega_0 t. \tag{5}$$

The amplitude a of the signal is determined by the amplification in the repeater and the downlink losses in the link following the limiter.

At the receiver the signal $g(t)$ is multiplied by the correctly phased reference $2 \cos \omega_0 t$ and then low-pass filtered to remove the double-frequency components. The filter output is given by

$$z(t) = a \cos \phi(t) + u(t). \tag{6}$$

B. Calculation of Error Probability

At the receiver output, the decision as to whether a mark or a space was transmitted is made on the basis of whether the filter output at the sampling instant is positive or negative. In order to calculate the probability of error of the decision, we need to determine the probability density function $q(z)$ of the filter output. This is obtained directly from (6) as the convolution of the probability density function $p(\phi)$ of the phase and a Gaussian density function for the inphase downlink noise. The result is

$$q(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^{2\pi} \exp \left(\frac{-(z - a \cos \phi)^2}{2\sigma^2} \right) p(\phi) \, d\phi \tag{7}$$

where σ^2 is the noise power at the receiver input. This integral representation for $q(z)$ will be convenient for the calculation of the error probability.

We assume that each transmitted symbol is statistically independent of its predecessors and that mark and space signals are equally probable. Hence, since the noise at the low-pass filter output has a zero-mean value, the probability of error P_e of a decision is equal to the error probability for either a mark or a space signal. Thus

$$P_e = P_{em} = P_{es}. \tag{8}$$

The probability of error in detecting a mark ($\theta = 0$) is equal to the probability that the filter output will be negative at the sampling instant, viz.,

$$P_e = \int_{-\infty}^{0} q(z) \, dz. \tag{9}$$

Substituting the expression for $q(z)$ in (9) and changing the order of integration yields

$$P_e = \int_0^{2\pi} p(\phi) \left[\frac{1}{\sqrt{2\pi\sigma^2}} \right.$$
$$\left. \cdot \int_{-\infty}^{0} \exp \left(\frac{-(z - a \cos \phi)^2}{2\sigma^2} \right) dz \right] d\phi. \tag{10}$$

The integral in the brackets can be expressed in terms of the error function [5]. Thus (10) may be written as

$$P_e = \frac{1}{2} \int_0^{2\pi} \left[1 - \mathrm{erf} \left(\frac{a}{\sqrt{2} \, \sigma} \cos \phi \right) \right] p(\phi) \, d\phi$$

$$= \frac{1}{2} - \frac{1}{2} \int_0^{2\pi} \mathrm{erf} \left(\frac{a}{\sqrt{2} \, \sigma} \cos \phi \right) p(\phi) \, d\phi. \tag{11}$$

Clearly, with no noise on the uplink, the error probability is unaffected by the presence of the limiter and is simply $P_e = \frac{1}{2} - \frac{1}{2} \mathrm{erf} (a/\sqrt{2} \, \sigma)$. Likewise, with no noise in the downlink the limiter will have no effect on P_e. This is easily seen by expressing $\cos \phi(t)$ in (6) as $[A + x(t)]/R(t)$, and noticing that P_e is the same as the probability of the inphase component $x(t)$ of the uplink noise exceeding the signal amplitude A, which is simply $P_e = \frac{1}{2} - \frac{1}{2} \mathrm{erf} (A/\sqrt{2} \, \sigma_1)$.

To evaluate the integral in (11), we use the representation for the error function [6]

$$\text{erf}\left(\frac{a}{\sqrt{2}\,\sigma}\cos\phi\right) = \frac{2}{\pi}\int_0^\infty \frac{\sin(va\cos\phi)}{v}\exp\left(\frac{-v^2\sigma^2}{2}\right)dv \quad (12)$$

which is seen by Plancherel's theorem to be correct. This representation is convenient for expressing the error function in the form of a Fourier series. Using the expression

$$\sin(va\cos\phi) = 2\sum_{n=0}^\infty (-1)^n J_{2n+1}(va)\cos(2n+1)\phi \quad (13)$$

we find that (12) becomes

$$\text{erf}\left(\frac{a}{\sqrt{2}\,\sigma}\cos\phi\right) = \frac{4}{\pi}\sum_{n=0}^\infty (-1)^n \int_0^\infty \frac{J_{2n+1}(va)}{v}$$
$$\cdot \exp\left(\frac{-v^2\sigma^2}{2}\right)dv\,\cos(2n+1)\phi. \quad (14)$$

The integral here is attributed to Weber and Sonine, and its solution is given in terms of confluent hypergeometric series [7]. The result is

$$\text{erf}\left(\frac{a}{\sqrt{2}\,\sigma}\cos\phi\right)$$
$$= \frac{2}{\pi}\sum_{n=0}^\infty \frac{(-1)^n\Gamma(n+\frac{1}{2})}{(2n+1)!}\left(\frac{a}{\sqrt{2}\,\sigma}\right)^{2n+1}$$
$$\cdot {}_1F_1\left(n+\frac{1}{2},\,2n+2,\,-\frac{a^2}{2\sigma^2}\right)\cos(2n+1)\phi \quad (15)$$

where ${}_1F_1(\cdot)$ is the confluent hypergeometric series and $\Gamma(\cdot)$ is the gamma function. Substituting (15) in (11) yields

$$P_e = \frac{1}{2} - \frac{1}{\pi}\sum_{n=0}^\infty \frac{(-1)^n\cdot\Gamma(n+\frac{1}{2})}{(2n+1)!}\left(\frac{a}{\sqrt{2}\,\sigma}\right)^{2n+1}$$
$$\cdot {}_1F_1\left(n+\frac{1}{2},\,2n+2,\,-\frac{a^2}{2\sigma^2}\right)$$
$$\cdot \int_0^{2\pi}\cos(2n+1)\phi\cdot p(\phi)\,d\phi. \quad (16)$$

A number of equivalent representations for the probability density function $p(\phi)$ are known in the literature [8], [9]. They would lead to equivalent but different expressions for the error probability. In this paper we shall consider two different representations for $p(\phi)$ and derive the corresponding expressions for the error rate. First, we consider the integral representation for $p(\phi)$. When a mark is transmitted, i.e., $\theta(t) = 0$, $p(\phi)$ is given by [9]

$$p(\phi) = \frac{\exp(-A^2/2\sigma_1^2)}{2\pi\sigma_1^2}$$
$$\cdot \int_0^\infty R\exp\left[\frac{-(R^2-2AR\cos\phi)}{2\sigma_1^2}\right]dR. \quad (17)$$

The integral in (16) may now be evaluated with the aid of (17). The result is [10]

$$\int_0^{2\pi}\cos(2n+1)\phi p(\phi)\,d\phi$$
$$= \left(\frac{A}{\sqrt{2}\,\sigma_1}\right)^{2n+1}$$
$$\cdot \frac{\Gamma(n+\frac{3}{2})}{(2n+1)!}\,{}_1F_1\left(n+\frac{1}{2},\,2n+2,\,-\frac{A^2}{2\sigma_1^2}\right). \quad (18)$$

Substituting (18) in (16), we obtain the probability of error

$$P_e = \frac{1}{2} - \frac{1}{\pi}\sum_{n=0}^\infty \frac{(-1)^n\cdot\Gamma(n+\frac{1}{2})\cdot\Gamma(n+\frac{3}{2})}{[(2n+1)!]^2}\rho_1^{2n+1}\cdot\rho_2^{2n+1}$$
$$\cdot {}_1F_1(n+\tfrac{1}{2},\,2n+2,\,-\rho_1^2)$$
$$\cdot {}_1F_1(n+\tfrac{1}{2},\,2n+2,\,-\rho_2^2) \quad (19)$$

where $\rho_1^2 = A^2/2\sigma_1^2 = $ limiter input or uplink signal-to-noise ratio, and $\rho_2^2 = a^2/2\sigma^2 = P_t/\sigma^2 = $ receiver input or downlink power-to-noise ratio. Note that P_t contains both signal and retransmitted limiter input noise power. We now show that the series in (19) is convergent. Since

$$|{}_1F_1(a,b,-x)| \le 1, \quad \text{for } b \ge a > 0 \le x \quad (20)$$

the series in (19) will converge if the dominating series obtained by setting both confluent hypergeometric functions equal to one converges. The resulting series is absolutely convergent for all values of ρ_1^2 and ρ_2^2, as can be shown easily by d'Alembert's ratio test for absolute convergence

$$\lim_{n\to\infty}\left|\frac{U_{n+1}}{U_n}\right| = \lim_{n\to\infty}\left|\frac{(n+\frac{1}{2})(n+\frac{3}{2})}{(2n+3)^2\cdot(2n+2)^2}(\rho_1\rho_2)^2\right| = 0. \quad (21)$$

An alternative expression for P_e is obtained by using the relationship

$${}_1F_1(n+\tfrac{1}{2},\,2n+2,\,-\rho^2)$$
$$= \frac{\exp(-\rho^2/2)\cdot n!\cdot 2^{2n}}{\rho^{2n}}\left[I_n\left(\frac{\rho^2}{2}\right)+I_{n+1}\left(\frac{\rho^2}{2}\right)\right] \quad (22)$$

where $I_n(\cdot)$ is the modified Bessel function.

The expression for P_e then becomes

$$P_e = \frac{1}{2} - \frac{\rho_1\cdot\rho_2}{2}\exp\left[\frac{-(\rho_1^2+\rho_2^2)}{2}\right]\sum_{n=0}^\infty \frac{(-1)^n}{2n+1}$$
$$\cdot \left\{I_n\left(\frac{\rho_1^2}{2}\right)+I_{n+1}\left(\frac{\rho_1^2}{2}\right)\right\}$$
$$\cdot \left\{I_n\left(\frac{\rho_2^2}{2}\right)+I_{n+1}\left(\frac{\rho_2^2}{2}\right)\right\}. \quad (23)$$

Next we derive an expression for the error rate by using a closed-form representation for the probability density function of the phase [9]

$$p(\phi) = \frac{\exp(-\rho_1^2)}{2\pi} + \sqrt{\frac{\rho_1^2}{4\pi}}\cos\phi\exp(-\rho_1^2\sin^2\phi)$$
$$\cdot [1+\text{erf}(\rho_1\cos\phi)]. \quad (24)$$

Substituting (24) in (11) and performing the integration provides an expression for the error probability in terms of Rice's Ie function [11]

$$P_e = \tfrac{1}{2}[1 - \sqrt{1 - k^2} \, Ie(k,x)] \qquad (25)$$

where

$$k = \frac{\rho_1{}^2 - \rho_2{}^2}{\rho_1{}^2 + \rho_2{}^2}, \qquad x = \frac{\rho_1{}^2 + \rho_2{}^2}{2} \qquad (26)$$

and

$$Ie(k,x) = \int_0^x \exp(-t) \cdot I_0(kt) \, dt. \qquad (27)$$

When $\rho_1 = \rho_2$, $Ie(0,x) = 1 - \exp(-x)$, and the error probability is then simply

$$P_e = \tfrac{1}{2} \exp(-\rho_1{}^2). \qquad (28)$$

C. Asymptotic Values for the Error Probability

It is interesting to examine the behavior of P_e for small and large SNRs. An asymptotic result for ρ_1 and ρ_2 much less than unity is obtained by retaining only the first term of the infinite series in (19) and the first term of the confluent hypergeometric series appearing in it

$$P_e \approx \tfrac{1}{2}(1 - \rho_1\rho_2), \qquad \text{for } \rho_1\rho_2 \ll 1. \qquad (29)$$

When both $\rho_1{}^2$ as well as $\rho_1{}^2 - \rho_2{}^2$ are large, we can find an asymptotic expression for P_e by using the expansion of $Ie(k,x)$ for large x [11]

$$Ie(k,\infty) \approx \frac{1}{\sqrt{1 - k^2}} - \left[\frac{1}{\sqrt{2k(1 - k)}} \left(1 - \operatorname{erf} \sqrt{x(1 - k)} \right) \right] \qquad (30)$$

in (25). We thus obtain

$$P_e = \frac{1}{2} \cdot \frac{\rho_1}{\sqrt{\rho_1{}^2 - \rho_2{}^2}} [1 - \operatorname{erf} \rho_2] \qquad (31)$$

$$\approx \frac{\rho_1}{2\rho_2} \cdot \frac{\exp(-\rho_2{}^2)}{\sqrt{\pi(\rho_1{}^2 - \rho_2{}^2)}}, \qquad \text{for } \rho_1 > \rho_2 \gg 1$$

$$\text{and } \rho_1{}^2 - \rho_2{}^2 \gg 1. \qquad (32)$$

When $\rho_1{}^2$ is very much larger than $\rho_2{}^2$, (31) reduces to

$$P_e = \tfrac{1}{2}[1 - \operatorname{erf} \rho_2], \qquad \text{for } \rho_1 \to \infty \qquad (33)$$

which is the expected result for ideal limiting in the absence of uplink noise. Because P_e is a symmetric function of $\rho_1{}^2$ and $\rho_2{}^2$, the corresponding results for $\rho_2{}^2 > \rho_1{}^2$ are obtained by simply interchanging $\rho_1{}^2$ and $\rho_2{}^2$ in (31)–(33).

III. Numerical Results

The results of numerical evaluation of the error rate are shown in Fig. 2 as a function of $\rho_1{}^2$ for various values of $\rho_2{}^2$. Similar curves for P_e as a function of $\rho_2{}^2$ with $\rho_1{}^2$ as the parameter are obtained by merely interchanging the labels $\rho_1{}^2$ and $\rho_2{}^2$ in Fig. 2, since the expression for the error rate is symmetric in $\rho_1{}^2$ and $\rho_2{}^2$. Both (19) and (23) were programmed, and the results obtained were identical,
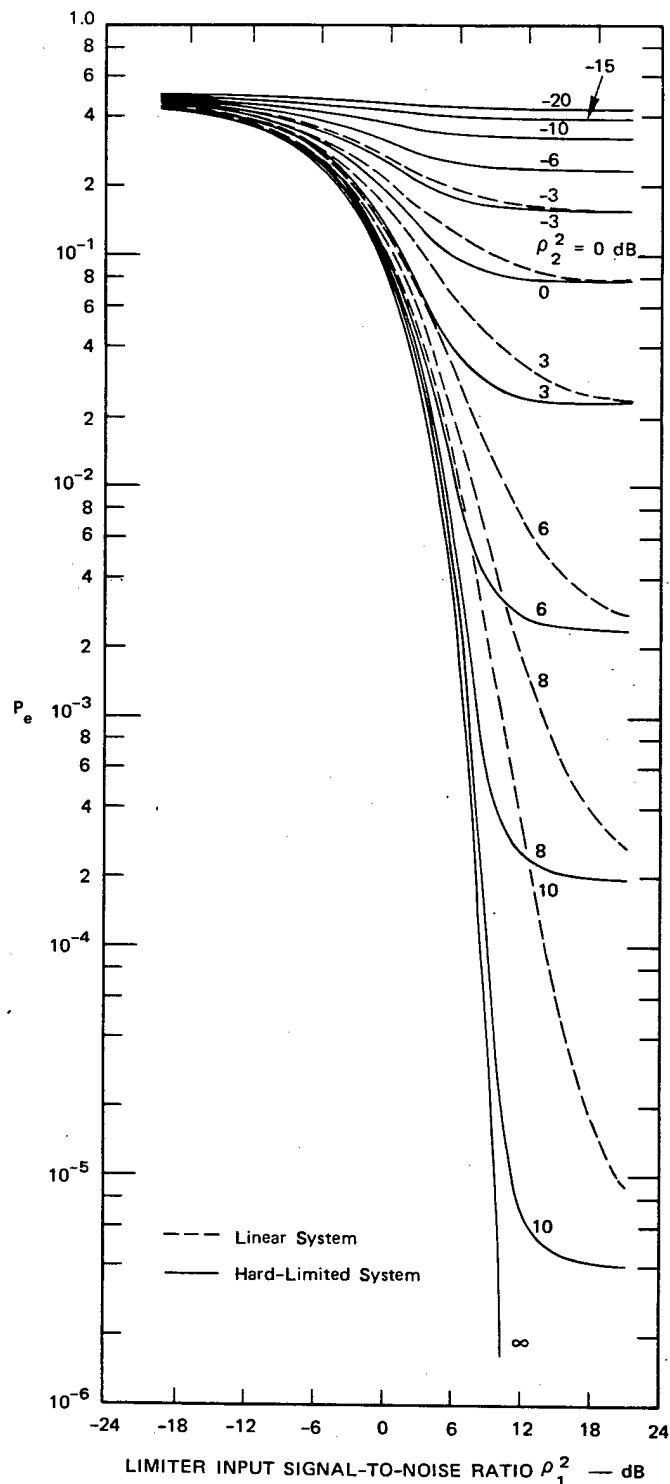
Fig. 2. PSK error rate as function of limiter input signal-to-noise ratio for constant values of received power-to-noise ratio.

as expected. Equation (25) was not used, because we had not yet discovered it. When both ρ_1^2 and ρ_2^2 are large, computation of the error rate from either (19) or (23) becomes difficult, and it is important that the series in (19) and (23) be computed very accurately. We found that (23) was easier to use than (19) at low error rates because the computation of the Bessel functions in (23) requires fewer terms of their power series expansion than do the two hypergeometric functions appearing in (19). The results obtained by using the asymptotic expression (32) for P_e when ρ_1^2 and ρ_2^2 are both large were found to be in excellent agreement with the numerical evaluation of (23).

The computation of the error rate from (25) would require accurate calculation of the Ie function defined by (27). Rice's table [11] for the Ie function does not give sufficient coverage for calculation of error rates under conditions of practical interest. The Ie function can be calculated by either numerical integration of (27) or from the series given by Rice [11]

$$Ie(k,x) = \sum_{n=0}^{\infty} (k/2)^{2n} \frac{(2n)!}{n!\,n!} A_n$$

where

$$A_n = 1 - \left[1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{2n}}{(2n)!} \right] \exp(-x).$$

A. Comparison with Linear-Repeater Performance

The dashed lines in Fig. 2 show the error rate for a linear system, i.e., a system with the limiter replaced by a linear repeater having the same average output power as the limiting repeater in Fig. 1. The error probability for a linear system, which has a Gaussian noise distribution at the receiver output, is

$$P_e = \tfrac{1}{2}[1 - \operatorname{erf}(\rho/\sqrt{2})] \tag{34}$$

where ρ^2 is the receiver output SNR. The repeater output power $P_t = a^2/2$ is divided between the signal and the uplink noise in proportion to their power levels at the repeater input. At the receiver the received waveform is processed by using correlation detection with baseband zonal filtering and sampling. The correlation operation suppresses the quadrature components of both retransmitted uplink and downlink noise, thereby providing a 3-dB improvement in the SNR at the receiver output over the SNR at its input. The output SNR for the linear system is thus

$$\rho^2 = 2 \frac{\rho_1^2 \rho_2^2}{1 + \rho_1^2 + \rho_2^2}. \tag{35}$$

Using two terms of the power-series expansion of the error function in (34) and approximating (35) as $\rho = \sqrt{2}\,\rho_1\rho_2$, we find that the error probability at low SNRs with a linear repeater is approximately

$$P_e \approx \frac{1}{2}\left(1 - 2\frac{\rho_1\rho_2}{\sqrt{\pi}}\right). \tag{36}$$

For a hard-limiting repeater the corresponding expression is given by (29). Thus we see that at low SNRs the linear repeater yields better performance than the hard limiter (although both do badly). The latter requires a 1.05-dB increase in the product of the two SNRs to obtain the same error probability.

Fig. 2 shows a bottoming of the error rate for both a linear and a hard-limiting system representing an irreducible error probability that depends on the noise present either at the repeater or the receiver front end, depending upon whether the abscissa is ρ_2^2 or ρ_1^2. It is seen in Fig. 2 that, as ρ_1^2 tends to infinity, both systems tend to the same limit given by (33). The significant difference, however, is that a hard-limiting system approaches the limit much faster, i.e., at lower values of ρ_1^2, than a linear system for any given value of ρ_2^2.

At large values of ρ_1^2 and ρ_2^2 the error rate for the linear system is obtained by using the first term of the asymptotic expansion for the error function in (34)

$$P_e \approx \frac{1}{2}\sqrt{\frac{\rho_1^2 + \rho_2^2}{\pi\rho_1^2\rho_2^2}} \exp\left(-\frac{\rho_1^2\rho_2^2}{\rho_1^2 + \rho_2^2}\right),$$

$$\text{for large } \rho_1 \text{ and } \rho_2. \tag{37}$$

The dominant factor here is the exponential, while in the case of the limiter with $\rho_1 \geq \rho_2$ it is $\exp(-\rho_2^2)$ as in (28) and (32). Comparing the respective exponents, we see that the limiter affords an SNR advantage of $1 + \rho_2^2/\rho_1^2 \leq 2$, i.e., at most 3 dB in each of the two links.

If the larger of the two SNRs, ρ_1^2, is increased while the smaller one, ρ_2^2, is held fixed, then we would obtain the same exponent for the linear system in (37) as with the limiter when ρ_1^2 is infinite. In this sense the limiter exhibits an unbounded SNR advantage. For example, we see from Fig. 2 that for $P_e = 2.4 \times 10^{-4}$ with $\rho_2^2 = 8$ dB there is a 9-dB horizontal spread in the value of ρ_1^2 between the solid and dashed curves. However, just increasing ρ_2^2 to 10 dB for the linear system suffices to obtain the same P_e with the same $\rho_1^2 = 12.5$ dB, and so the limiter affords only a 2-dB advantage in this sense. In fact, whenever the larger SNR exceeds the smaller by more than 3 dB, it would suffice to increase just the smaller one for the linear system by less than 3 dB. When the larger SNR exceeds the smaller by less than 3 dB, both SNRs must be increased to 3 dB more than the smaller.

Springett and Simon [12] have shown that for a hard-limiting system, in the absence of downlink noise, a large SNR improvement $\rho^2 \approx 8\,\rho_1^4$ is achieved at large values of ρ_1^2. Unfortunately this improvement in ρ^2 has no bearing on the error rate.

IV. OPTIMUM NONLINEARITY

So far we have supposed that the nonlinearity is a hard limiter, delivering a sine wave of amplitude a to the receiver regardless of its input amplitude R. More generally, a will be a function of R, say $\sigma h(AR/\sigma_1^2)$, the constant factors σ and A/σ_1^2 being included here to simplify the form of our

analysis. To take account of this nonlinearity it is only necessary to replace a in (11) by $\sigma h(AR/\sigma_1^2)$. Using (17) for $p(\phi)$ in (11) and substituting $r = AR/\sigma_1^2$, we thus get

$$P_e = \frac{1}{2} - \frac{\exp(-\rho_1^2)}{8\pi\rho_1^2} \int_0^{2\pi} \int_0^\infty r$$
$$\cdot \exp\left[\frac{-(r^2 - 4r\rho_1^2 \cos \phi)}{4\rho_1^2}\right] \operatorname{erf}\left(\frac{h(r)}{\sqrt{2}} \cos \phi\right) dr\, d\phi. \tag{38}$$

Since r has a Rice–Nakagami distribution, the average repeater power delivered to the receiver is the mean value of $\frac{1}{2}\sigma^2 h^2(r)$

$$P_t = \frac{\sigma^2}{4\rho_1^2} \exp(-\rho_1^2) \int_0^\infty r \exp\left(-\frac{r^2}{4\rho_1^2}\right) I_0(r) h^2(r)\, dr. \tag{39}$$

We now seek that $h(r)$ that minimizes (38) under the condition that (39) has a prescribed value. Treating the integrals with respect to r as summations over discrete values of $h(r)$, we solve this problem by simply differentiating (38) and (39) with respect to the value taken by h at one specific value of r. Introducing the Lagrange multiplier $-1/(\sqrt{2\pi}\, \sigma^2 L)$ for the derivative of (39), setting the sum of this plus the derivative of (38) equal to zero and cancelling out the common factor

$$\frac{\exp(-\rho_1^2)}{4\pi\sqrt{2\pi}} \exp\left(-\frac{r^2}{4\rho_1^2}\right) dr$$

we thus obtain [13]

$$I_0(r)h(r)$$
$$= \frac{L}{2\pi} \int_0^{2\pi} \exp(r \cos \phi) \exp\left(-\tfrac{1}{2}h^2(r) \cos^2 \phi\right) \cos \phi\, d\phi$$
$$= \frac{2L}{\pi} \int_0^{\pi/2} \sinh(r \cos \phi) \exp\left(-\tfrac{1}{2}h^2(r) \cos^2 \phi\right) \cos \phi\, d\phi. \tag{40}$$

Functions $h(r)$ satisfying (40) can be found by solving for L and determining L as a function of h and r. The contours over the (h,r) plane along which L is constant describe the required functions.

To determine what constant value L must have, the resulting $h(r)$, which depends on L, is substituted into (39), and L is found as a function of ρ_1^2 and $P_t/\sigma^2 = \rho_2^2$. From (40) we see that $h(r) = 0$ when $L = 0$ and that $h(r)$ is an unbounded, increasing function of L for every r. It follows from (39) that $L = 0$ when $\rho_2^2 = 0$ and that L is an unbounded, increasing function of ρ_2^2.

Although the solution of the integral equation (40) for $h(r)$ is not a straightforward matter, it can be solved analytically in extreme cases. For example, when L is small, h too is small, and the second exponential in the first line of (40) can be replaced by unity, giving us

$$h(r) = LI_1(r)/I_0(r). \tag{41}$$

This function increases linearly from zero for small values of r, and it asymptotically rises toward the value L for large r; i.e., (41) represents a sort of soft limiter, and the larger the uplink signal-to-noise ratio, the harder this limiter becomes on account of the factor A/σ_1^2 in the argument of h. It is remarkable that (41) is the same nonlinearity [13] that maximizes the output signal-to-noise ratio of the repeater output.

More generally, we see from (40) that $h(0)$ is always zero, that

$$h(r) = \frac{1}{2} Lr - \frac{3L^3 + 4L}{64} r^3 + \cdots$$

and that, as $r \to \infty$, $h(r)$ approaches the value satisfying the equation

$$h(\infty) = L \exp\left[-\tfrac{1}{2}h^2(\infty)\right] \tag{42}$$

as the integrand of (40) then peaks at $\phi = 0$ and the right-hand side is asymptotically

$$L \exp\left[-\tfrac{1}{2}h^2(\infty)\right] \cdot I_1(r).$$

Thus, whenever the uplink signal-to-noise ratio is very large, a hard limiter is the optimum nonlinearity. In this case we can determine L by writing (39) in the form $\rho_2^2 = \tfrac{1}{2}h^2(\infty)$ and substituting (42). Hence, for large ρ_1^2 we have

$$L = \sqrt{2}\, \rho_2 \exp(\rho_2^2). \tag{43}$$

When L is large, we have

$$h(r) \approx \sqrt[6]{\frac{2}{\pi}} \cdot \sqrt[3]{\frac{Lr}{I_0(r)}} \tag{44}$$

provided that (44) is large in comparison with $r + 1$. To see this notice that a large L implies a large $h(r)$, which causes the integrand of (40) to be negligible except near $\phi = \tfrac{1}{2}\pi$, where $\cos \phi \approx \tfrac{1}{2}\pi - \phi$. For large downlink signal-to-noise ratios, therefore, the optimum nonlinearity (44) rises steeply from zero for small r, reaches a maximum where $r = 1.61$, and then falls toward the value satisfying (42). This is, in fact, qualitatively the same as the output-versus-input characteristic of a traveling-wave tube.

To see that the $h(r)$ determined by (40) must yield a minimum rather than a maximum error probability, we notice that P_e can be made arbitrarily close to unity simply by putting $h(r) = 0$, for all r, except in some interval, which is allowed to shrink to zero width. Because such an $h(r)$ becomes a kind of delta function rather than a proper function, it fails to show up in our solution, and so we obtain only the optimum $h(r)$ from (40) and not the pessimum.

V. Conclusions

General expressions have been derived for the probability of error in detecting a binary phase-modulated sinusoidal signal after transmission through a hard-limiting channel. The receiver employs cross correlation followed by zonal low-pass filtering for the detection of the signal. The results are compared with those for a linear channel and a

correlation receiver not using postdetection integration. The error rate is found to be smaller than that of a linear channel for the range of SNRs that will be of interest in practice. In the absence of either repeater or receiver input noise the error-rate expression is identical with that for a linear channel with no postdetection integration.

The results of this paper can also be extended as in Lyons [3] to include the effect of postdetection integration via multiple sampling and majority decision. This probability of error is then given by the binomial distribution into which we need only to substitute the error probability for a single sample that is obtained in this paper. The result is a monotonic function of the error probability for a single sample, and therefore the conclusions reached in this paper concerning the helpfulness of limiting with a single sample detection also apply with postdetection integration.

The optimum nonlinearity has been determined and turns out to be a limiter whenever the uplink SNR is large. Remarkably, this optimum nonlinearity behaves like a traveling-wave tube when the downlink SNR is large, with its output falling slowly as the input level increases beyond a certain value. It may be mentioned that limiting can similarly be shown to reduce the error probability (although to a smaller extent) for quadriphase and higher-order phase-shift-keying.

REFERENCES

[1] I. Jacobs, "The effects of video clipping on the performance of an active satellite PSK communication system," *IEEE Trans. Commun. Technol.*, vol. COM-13, pp. 195–201, June 1965.

[2] L. D. Davisson and L. B. Milstein, "On the performance of digital communication systems with bandpass limiters—Part I: One-link systems and —Part II: Two-link systems," *IEEE Trans. Commun.*, vol. COM-20, pp. 972–980, Oct. 1972.

[3] R. G. Lyons, "The effect of bandpass nonlinearity on signal detectability," *IEEE Trans. Commun.*, vol. COM-21, pp. 51–60, Jan. 1973.

[4] P. C. Jain, "Performance of hard-limited spread-spectrum transmission systems," Rome Air Development Center, Griffiss Air Force Base, N.Y., Contract F30602-71-C-0338.

[5] M. Schwartz, *Information Transmission, Modulation, and Noise.* New York: McGraw-Hill, 1959, p. 373.

[6] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics.* New York: Springer, 1966, p. 418.

[7] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Dover, 1965, p. 486.

[8] D. Middleton, *An Introduction to Statistical Communication Theory.* New York: McGraw-Hill, 1960, pp. 416–418.

[9] M. Schwartz, *Information Transmission, Modulation, and Noise.* New York: McGraw-Hill, 1959, p. 411.

[10] J. L. Lawson and G. E. Uhlenbeck, *Threshold Signals.* New York: McGraw-Hill, 1950, p. 173.

[11] S. O. Rice, "Statistical properties of a sine wave plus random noise," *Bell Syst. Tech. J.*, vol. 27, pp. 109–157, Jan. 1948.

[12] J. C. Springett and M. K. Simon, "An analysis of the phase coherent–incoherent output of the bandpass limiter," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 42–49, Feb. 1971.

[13] N. M. Blachman, "Detectors, bandpass nonlinearities, and their optimization: inversion of the Chebyshev transform," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 398–404, July 1971.